



VYSOKÉ UČENÍ TECHNICKÉ V BRNĚ
BRNO UNIVERSITY OF TECHNOLOGY



FAKULTA PODNIKATELSKÁ
ÚSTAV INFORMATIKY

FACULTY OF BUSINESS AND MANAGEMENT
INSTITUT OF INFORMATICS

ZÁLOHOVÁNÍ DAT A DATOVÁ ÚLOŽIŠTĚ

DATA BACKUP AND STORAGE

BAKALÁŘSKÁ PRÁCE

BACHELOR'S THESIS

AUTOR PRÁCE

AUTHOR

ONDŘEJ CABALKA

VEDOUCÍ PRÁCE

SUPERVISOR

ING. JIŘÍ KŘÍŽ, PH.D.

BRNO 2011

ZADÁNÍ BAKALÁŘSKÉ PRÁCE

Cabalka Ondřej

Manažerská informatika (6209R021)

Ředitel ústavu Vám v souladu se zákonem č.111/1998 o vysokých školách, Studijním a zkušebním řádem VUT v Brně a Směrnicí děkana pro realizaci bakalářských a magisterských studijních programů zadává bakalářskou práci s názvem:

Zálohování dat a datová úložiště

v anglickém jazyce:

Data Backup and Storage

Pokyny pro vypracování:

Úvod

Vymezení problému a cíle práce

Teoretická východiska práce

Analýza problému a současné situace

Vlastní návrhy řešení, přínos návrhů řešení

Závěr

Seznam použité literatury

Přílohy

Seznam odborné literatury:

- BASL, J.; BLAŽÍČEK, R. Podnikové informační systémy. Praha: Grada Publishing, 2007. 288 s. ISBN 978-80-247-2279-5.
- DOSEDĚL, T. Počítačová bezpečnost a ochrana dat. Brno : Computer Press, 2004. 190 s. ISBN 80-251-0106-1.
- CHARVÁT J. Firemní strategie pro praxi. Praha: Grada Publishing, 2006. 204 s. ISBN 80-247-1389-6.
- SODOMKA, P. Informační systémy v podnikové praxi. Brno: Computer Press, 2006. 352 s. ISBN 80-251-1200-4.
- TVRDÍKOVÁ, M. Aplikace moderních informačních technologií v řízení firmy. Praha: Grada Publishing, 2008. 176 s. ISBN 978-80-247-2728-8.
- VOŘÍŠEK J. Strategické řízení informačního systému a systémová integrace. Praha : Management Press, 2006. 324 s. ISBN 80-85943-40-9.

Vedoucí bakalářské práce: Ing. Jiří Kříž, Ph.D.

Termín odevzdání bakalářské práce je stanoven časovým plánem akademického roku 2010/2011.

L.S.

Ing. Jiří Kříž, Ph.D.
Ředitel ústavu

doc. RNDr. Anna Putnová, Ph.D., MBA
Děkan fakulty

V Brně, dne 02.06.2011

ABSTRAKT

Tato bakalářská práce upozorňuje na hojně opomíjenou důležitost firemních dat a jejich ochranu. Seznamuje s procesem zálohování jako základním prvkem bezpečnosti a efektivnosti informačního systému. Součástí práce jsou základní i pokročilé metody zálohování vhodné pro malé a střední podniky, v závislosti na jejich potřebě a finančních možnostech.

KLÍČOVÁ SLOVA

data, ztráta dat, zálohování, obnova, archivace, pevný disk

ABSTRACT

This work highlights the frequently overlooked importance of corporate data and their protection. It introduces the backup process as an essential element of safety and efficiency of information system. The work is both basic and advanced backup methods suitable for small and medium enterprises, depending on their needs and budget.

KEY WORDS

data, data loss, backup, restore, archiving, hard disk

BIBLIOGRAFICKÁ CITACE

CABALKA, O. *Zálohování dat a datová úložiště*. Brno : Vysoké učení technické v Brně, Fakulta podnikatelská, 2011. 57 s. Vedoucí bakalářské práce Ing. Jiří Kříž, Ph.D.

ČESTNÉ PROHLÁŠENÍ

Prohlašuji, že předložená bakalářská práce je původní a zpracoval jsem ji samostatně. Prohlašuji, že citace použitých pramenů je úplná, že jsem v práci neporušil autorská práva (ve smyslu zákona č. 121/2000 Sb. O právu autorském a o právech souvisejících s právem autorským).

V Brně dne

.....
podpis

PODĚKOVÁNÍ

Rád bych tímto poděkoval Ing. Jiřímu Křížovi, Ph. D. a Ing. Josefu Veškrnovi za poskytnutí mnoha užitečných rad a poznatků, bez kterých by tato bakalářská práce jen těžko vznikla.

OBSAH

Úvod.....	9
1 VYMEZENÍ PROBLÉMU A CÍLE PRÁCE.....	10
2 TEORETICKÁ VÝCHODISKA PRÁCE	11
2.1 Rozdělení záznamu elektronických dat podle technologie	11
2.2 Životnost prostředků pro ukládání dat	13
2.3 Trvanlivost uložených dat.....	15
2.4 Porovnání rizik s ohledem na technologie záznamu	16
2.5 Další technologie.....	19
2.6 Doporučený model řízené archivace na dobu stovek let.....	21
2.7 Technologie RAID obecně.....	24
2.8 Cloud computing.....	26
2.9 Amazon Web Services	27
3 ANALÝZA PROBLÉMU A SOUČASNÁ SITUACE	28
3.1 Vybraný případ pro provedení zálohování dat.....	28
3.2 Současný stav	28
3.3 Důvody nevyhovujícího systému	34
4 VLASTNÍ NÁVRHY ŘEŠENÍ, PŘÍNOS NÁVRHU ŘEŠENÍ	37
4.1 Možnosti řešení	37
4.2 Vlastní zálohování ve firmě	41
4.3 Externí zálohování	46
4.4 Pohled do budoucnosti	50
4.5 Doporučení jedné z navržených možností	51
4.6 Přínos návrhu řešení.....	53
Závěr	54
Seznam literatury	55
Seznam obrázků	57

ÚVOD

Když se člověk neznalý v oboru informačních technologií podívá na obyčejný pevný disk, jak je komplikovaný a zběžným pohledem těžko popsatelný objekt, který se po bližším zkoumání dá charakterizovat jako zařízení kombinující elektroniku a jemnou mechaniku, následně po delším dohledávání dostupných informací zjistí, že se zdaleka nejedná o jedinou technologii na uchovávání dat. Jsou tu virtuální knihovny, robotické knihovny, nebo moderní technologie mající zkratku **CLOUD**. Navíc velkým úskalím disků je náchylnost k poruše, což je zejména nepříjemné u serverů, kde cena uložených dat může dosahovat obrovských částek. Uživatel musí počítat s odstávkou serveru kvůli nutným opravám nebo s jeho nefunkčností při pravidelném zálohování, kdy se obnovují data, což má za následek značné prostoje mnoha uživatelů.

Proto jsem se rozhodl napsat tuto bakalářskou práci, abych detailněji porozuměl problematice zálohování a ukládání dat a ve spolupráci s firmou Agora plus, a.s., která je českou distribuční společností působící v oboru informačních technologií a též distribučním partnerem s přidanou hodnotou (Value Added Distributor) významných světových výrobců produktů z oblasti ukládání dat a serverových síťových řešení a technologií, najít vhodné řešení pro jejího zákazníka, který má zastarale fungující síť sloužící k úkladu dat. Firma Agora se mimo jiné zaměřuje na komplexní služby spojené s ukládáním a správou dat, tvorbu složitých podnikových **SAN** sítí neboli sítě k ukládání dat.

1 VYMEZENÍ PROBLÉMU A CÍLE PRÁCE

Cílem mé bakalářské práce je najít vhodné řešení, které nahradí stávající nevyhovující a zastaralý systém na zálohování a ukládání dat, pro zákazníka firmy Agora plus, a.s., aby dobře fungující firma mohla nadále bezproblémově pokračovat ve své činnosti bez zbytečných prodlev z důvodu nedostatečně ošetřeného systému zálohy dat.

Teoretická část obsahuje nejdříve informace o rozdělení záznamu elektronických dat podle technologie, kde se pozastavuji například u magnetického záznamu, optického záznamu, magneto-optického záznamu, polovodičové paměti a holografického záznamu. Poté se zaměřuji na životnost prostředků pro ukládání dat či jejich trvanlivost a další principy. Neméně důležité je objasnění pojmů v oblasti porovnání rizik s ohledem na technologie záznamu. K ukládání dat neodmyslitelně patří technologie robotické či virtuální knihovny, které zde proto zmiňuji. Nakonec v této části vysvětluji metodu zabezpečení dat proti selhání pevného disku technologií **RAID** a moderní službu **CLOUD** computing s možností externího uložení dat na serverech na Internetu pomocí specializovaného poskytovatele zaměřeného na tuto službu.

V další části stručně charakterizuji firmu, které chci pomoci, a proto musím zjistit její současný stav. Se současným stavem firmy se detailně seznamuji po prostudování materiálů firmy a následném sepsání prostředků **IT**, které firma využívá v současnosti. Na tuto kapitolu naváží důvody, proč stávající systém nevyhovuje požadavkům firmy, a v poslední čtvrté části nastiňuji dva vlastní návrhy řešení, které vybírám ze široké škály možností, posléze je porovnávám a na základě jejich výhod a nevýhod doporučuji co možná nejvhodnější variantu s danými přínosy.

2 TEORETICKÁ VÝCHODISKA PRÁCE

2.1 Rozdělení záznamu elektronických dat podle technologie

Magnetický záznam

Data jsou zaznamenávána do magnetické vrstvy materiálu změnou elektromagnetického pole zapisovací hlavy. Vznikají záznamové stopy soustředné - rotační záznam (pružný disk, pevný disk), nebo stopy podélné - lineární záznam (magnetické pásky s lineárním záznamem), případně podélné stopy s šikmým záznamem. Výhodou soustředného uspořádání stop je dostupnost dat po celém povrchu média v libovolném místě ve velmi krátkém časovém okamžiku.

Pokud neuvažujeme o technologii pružných disků, která je dnes na ústupu, používá se rotační záznam u magnetických pevných disků, které jsou navíc konstruovány jako uzavřené jednotky s velmi dobrou odolností proti negativním vlivům vnějšího prostředí. Nevýhodou pevných disků je jejich mechanická složitost, citlivost na otřesy a vyšší pořizovací cena.

Podélný záznam je využíván u záznamových médií typu magnetická páska. V některých případech může být nevýhodou sekvenční přístup při vyhledávání informací, při zápisu sekvenční přístup není na škodu. V mnoha případech je dosahováno efektivně stejné rychlosti zápisu jako u pevných disků. Sekvenční adresace datových bloků nedovoluje využívat kapacitu páskového média efektivně, smazané bloky nelze zaplnit novými informacemi. Proto je magnetická páska téměř výhradně používána jako cílové médium pro vytváření záloh nebo archivů. Při opakovaném zapisování jsou data buď přidávána za již stávající záznam, nebo je celá páska přepisována znovu od počátku. Doposud výraznou výhodou záznamu na magnetické pásky je její ekonomická efektivnost.

Podélný záznam v šikmých stopách postupně ztrácí svůj význam a zůstává doménou stále se zmenšujícího segmentu prostředků pro klasické jednoduché zálohování.

Optický záznam

Nejrozšířenějším druhem záznamu je tzv. lisování, kdy informace uložená na nosiči vzniká již při jeho výrobě a dále ji není možno měnit. Tento druh záznamu je využíván téměř výhradně k distribuci elektronických dat, v zábavním a multimediálním průmyslu. Cenová dostupnost je založena na značném množství vznikajících stejných kopií jednoho záznamu. Doposud nebyla aplikována cenově efektivní technologie lisování pro jednotlivé unikátní záznamy.

V oblasti informačních technologií je optický záznam použit jako zápis pomocí soustředěného paprsku světla určité vlnové délky do světločivého materiálu. Doposud se používá zápis na rotační médium, ale ne v soustředných stopách, nýbrž v jedné nebo více stopách spirálových. Přístup k informacím při čtení je potom podobný sekvenčnímu přístupu lineárních záznamů. Vzhledem k malému množství zapsaných datových bloků je však tento přístup poměrně rychlý.

Magneto-optický záznam

Jedná se o málo rozšířený způsob záznamu optickou cestou, který k vytváření jednotek informace v povrchu datového nosiče kromě světla současně využívá i působení magnetického pole. Uspořádání magneto-optického záznamu je kombinací záznamů v soustředných stopách a ve spirále - datové bloky jsou umístěny v různých hustotních zónách. Dochází tak k lepšímu využití kapacitních možností citlivého materiálu. Charakteristickou vlastností magneto-optických nosičů je jejich vysoká trvanlivost a odolnost proti vnějším vlivům.

Polovodičové paměti

Stále rostoucí kapacita rozličných druhů polovodičových pamětí a jejich nízké energetické nároky na udržení informací činí z této technologie pravděpodobně potenciálního nástupce dosavadních způsobů záznamu elektronických dat. Zásadní výhodou je velmi vysoká odolnost proti vlivům vnějšího prostředí a proti otřesům, jednoduchá dostupnost informací náhodnou adresací při čtení. Převládající nevýhodou je v dnešní době nízká kapacita a vysoká cena za jednotkovou kapacitu. Polovodičové

paměti představují velmi vážného kandidáta na nástupce klasických mechanických rotačních nebo lineárních druhů záznamů.

Holografický záznam

Z hlediska výtěžnosti záznamových materiálů se jeví jako nejperspektivnější optický holografický záznam, který řádově několikrát překračuje fyzikální možnosti stávajících magnetických nebo optických technologií. Vzhledem k využití optiky lze předpokládat i dosažení velmi vysokých rychlostí při zápisu a náhodný přístup k informacím při čtení. Holografický záznam může, ale v budoucnu nutně nemusí používat složité mechanické prvky. Představuje jeden z možných směrů dalšího vývoje technologií pro úschovu digitálních informací.

Mikromechanický záznam

V současné době existuje v laboratorním prostředí mechanicko-optický záznam, který díky nanotechnologiím dosahuje extrémně vysokých hustot záznamu a jsou u něj předpoklady i vysokých rychlostí čtení a zápisu. Vzhledem k použití velmi malých mechanických částí je tato technologie téměř odolná proti otřesům, ale zbývá vyřešit ještě celou řadu problémů se stabilitou záznamu a jeho odolností proti vnějším vlivům.

2.2 Životnost prostředků pro ukládání dat

V tomto oddílu budu posuzovat životnost faktickou i morální.

Jednotlivé druhy záznamu popsané v předešlé kapitole se vyznačují z fyzikálního hlediska různě dlouhou dobou, po kterou lze data bezpečně přečíst z citlivé záznamové vrstvy.

Magnetický záznam

Magnetický záznam ztrácí informace vlivem slábnutí elektrických sil, působících na elementární částice, které tvoří jednotku informace. Projektovaná doba trvanlivosti magnetického záznamu se pohybuje od 10 do 35 let.

Optický záznam

Optický záznam degraduje především působením vnějších vlivů a opotřebením při zacházení. Lisovaný optický záznam je v tomto ohledu nejtrvanlivější, ale pro běžnou archivaci se nepoužívá. Nejčastější příčinou ztráty informací na optických médiích je znehodnocení světločivé vrstvy chemickým působením okolního prostředí, které proniká do struktury vrstev konstrukce datového nosiče. Pro archivní účely se proto vybírají určité šarže pomocí měření kvality nebo se používají dražší výrobní postupy a materiály pro zajištění trvanlivosti. V automatizovaných archivech při nasazení mechanických manipulačních robotů se často používají ochranné kazety pro snížení pravděpodobnosti mechanického poškození povrchu a udržení citlivého média v optimálním prostředí bez prachu a světla. Projektovaná trvanlivost záznamu na optických médiích je od 10 do 100 let.

Magneto-optický záznam

Magneto-optický záznam byl již od počátku praktikován u médií určených k archivaci dat, proto u něj bývají výše uvedené konstrukční prvky uplatněny automaticky. Vždy je médium uloženo v pevné a uzavřené kazetě. Médium bývá často oboustranné pro dosažení vyšší kapacity. Robotické automaty navíc vyžadují horizontální rotaci ruky, aby bylo možno médium vkládat do čtecího zařízení v obou směrech. Vícehlavá čtecí zařízení, která by dokázala přistupovat na oba povrchy datového média, se neosvědčila pro svou mechanickou složitost a vysokou pořizovací cenu. Životnost magneto-optických, případně posledních generací optických médií je minimálně 50 let.

Polovodičové paměti

U polovodičových pamětí je životnost dána jen samotnou schopností média udržet záznam bez dodání vnější energie. To prozatím znemožňuje použít tato média pro archivní účely, protože trvanlivost záznamu je jen v jednotkách roků. Při zajištění trvalého napájení lze naopak považovat tento druh záznamu za nekonečný.

Další principy

U dalších nových principů ukládání informací zatím nelze jistě stanovit skutečnou dobu, po kterou si udrží zapsané informace.

Při vyhodnocování prostředků pro dlouhodobou archivaci dat není životnost záznamu rozhodujícím faktorem. Nastupuje hledisko morální životnosti, které bude probráno v následující kapitole.

2.3 Trvanlivost uložených dat

Archiv elektronických dat předpokládá, že informace v něm uložené musí přetrvávat v principu nekonečně dlouhou dobu beze změny. Z tohoto úhlu pohledu je životnost záznamu v technickém prostředku bezvýznamná. Mnohem důležitějším faktorem se stává plán řízení dostupnosti informací v dlouhodobém časovém horizontu. Pro účely archivace dat proto vyhoví téměř každé, v současnosti vyráběné médium, kde je trvanlivost projektována na 5 až 10 let.

Plán řízení dostupnosti informací proto zahrnuje minimálně tři principy fungování:

- A. Úschova technických i programových prostředků*
- B. Emulace původních programů na nových platformách*
- C. Trvalý program migrace prostředků pro úschovu dat*

Ad A)

Tato metoda je účinná jen po velmi krátkou dobu z hlediska celkových nároků na životnost dat. Udržování původních technických a programových prostředků se během času stává velmi nákladnou. Horizont použití této metody je 15 - 20 let.

Ad B)

Emulování především původních prostředí pro zajištění čitelnosti a zobrazitelnosti dat je jeden z nutných směrů, zvláště jedná-li se o multimediální data, dostupná nejen jako prosté informační jednotky, ale také jako ucelené programové

celky (prezentace, interaktivní programy, atp.). Existuje technologie tzv. univerzálních emulací, která se snaží řešit tento problém dlouhodobě a univerzálně.

Ad C)

Základní kámen života fungování elektronického archivu je program trvale udržitelné migrace dat. Cyklus generační výměny prostředků v informačních technologiích je prozatím velmi krátký a nic nenasvědčuje jeho prodlužování. Technologie záznamu pravidelně mění svoji kapacitu a výkon v souladu s postupným získáváním dalších poznatků z fyziky, chemie a mechaniky. Inovační cyklus se pohybuje v rozmezí 2 až 3 let. Za těchto podmínek je proto nutné projektovat digitální archiv s patřičnou mírou automatizace, která neslouží jen k manipulaci s informacemi a jejich nosiči, ale i k automatické migraci mezi přicházejícími generacemi.

Jednodušší případ migrace nastává v případě přesunu dat mezi prostředky stejného typu, ale s různou kapacitou. Složitější je migrace na zcela jiný typ záznamu. Oba dva případy se ale musí v plánu vyskytovat a jejich nasazení je nevyhnutelné.

Plán řízené dostupnosti musí pravděpodobně obsahovat všechny tři výše uvedené principy.

Dobře navržený dlouhodobý archiv by měl využívat i kombinací dvou různých technologií záznamu, jejichž inovační cyklus se pokud možno nepřekrývá. Migrace dat pak může probíhat pouze na jedné z platforem a nezatíží fungování archivu riziky přechodového období.

2.4 Porovnání rizik s ohledem na technologie záznamu

D. Pevné úložiště dat

E. Úložiště dat s výměnnými médii

Ad D)

Tento typ úschovy zahrnuje všechny druhy pevných disků a jejich uspořádání do diskových polí. Díky nedávnému vývoji v typech rozhraní a jeho dopadem na cenu a snadnost nasazení, se pevné disky **SATA** stávají důležitým prvkem velkokapacitních

zařízení pro hromadnou úschovu i archivaci dat. Uspořádání pevných disků do diskového pole s redundancí **RAID** zvyšuje spolehlivost a dostupnost dat. Nová generace disků s rozhraním **SATA** disponuje dostatečným výkonem i spolehlivostí (vybrané modelové řady jednotlivých výrobců).

Výhody:

- jednoznačně dominuje snadnost použití nevyžadující žádné zvláštní programy
- rychlost vyhledávání informací
- ochrana dat vytvářením záloh metadat (časové snímky)
- snadná migrace mezi generacemi
- zatím malá pravděpodobnost radikální změny technologie

Nevýhody:

- složité zálohování velkých objemů dat (soubory)
- velké množství malých souborů silně zatěžuje celý systém (souborový přístup)
- jednotková cena roste lineárně po dobu existence stejného typu disků
- formát záznamu na **HDD** (Hard Disk Drive) je zpravidla definován typem operačního systému a jeho souborovým systémem
- aplikace technologie **WORM** (vysvětleno níže) je možná, ale prakticky přináší významnou redukci výkonu
- prakticky nemožné manipulovat s celým klasickým pevným diskem jako s výměnným médiem

- ochrana dat dislokací je možná prakticky jen vzdáleným zrcadlením, což může být nákladné

Ad E)

Výměnná média, především magnetické pásky, jsou tradičním druhem prostředku pro archivaci. V minulosti byla volena zejména kvůli vysoké ekonomice provozu. Postupem času vznikla celá technologie archivace na výměnná média, včetně využití automatizovaných robotických knihoven. Existuje celá řada programů pro archivaci, často kombinujících archiv se zálohou. Programy nahrazují lidský faktor při manipulaci s médii a vedou podrobné záznamy o všech prováděných aktivitách.

Výhody:

- jednotková cena kapacity neroste lineárně, ale pomaleji
- je snadné vytvářet zálohy prostým klonováním existujících médií
- kazety lze snadno dislokovat na bezpečné místo
- formát dat na páskách je velmi obecný (tar) a dlouhodobě dostupný ve všech generacích operačních systémů. To je důležité z pohledu rekonstrukce dat v případech fatálních ztrát primárních dat a prostředků pro zápis/čtení
- blokový přenos dat na pásku je jednoduchý a velmi rychlý
- snadná implementace funkce **WORM**

Nevýhody:

- zranitelnost záznamu působením vnějších vlivů
- sekvenční přístup při čtení
- dlouhá inicializace média při založení do čtecí jednotky

- momentální cenové porovnání s pevnými disky **SATA** není příliš výhodné
- složitější migrace mezi generacemi, zejména při ukončení jedné etapy vývoje
- docela vysoká pravděpodobnost vymizení konkrétní technologie z trhu

2.5 Další technologie

WORM

Technologie **WORM** (Write Once Read Many) je často zmiňována v souvislosti s ukládáním dat trvalého charakteru, kde nežádoucí změna může znamenat znehodnocení nebo zneužití informace. Z fyzikálního hlediska jediným skutečným médiem **WORM** je optika (**CD-R** nebo **DVD-R** a jeho varianty). Původní **WORM** 800MB již dožívá v několika málo dedikovaných zařízeních a média již nejsou prakticky dostupná. Na optickém nosiči skutečně dojde k trvalé změně, kterou nelze stejným způsobem pozměnit. **WORM** aplikovaný na pevný disk nebo diskové pole představuje takovou úpravu mikrokódu zařízení, která zabraňuje modifikaci jednou uložených informací. Režie mikrokódu při zápisu je ale poměrně značná a ovlivňuje výkon zařízení. Navíc při poruše na úrovni mikrokódu může dojít ke zrušení **WORM** funkce a tím i k ohrožení původně trvalého záznamu. Podobně je tomu i u magnetických pásek, kde funkci **WORM** opět zajišťuje mikrokód čtecí/záznamové jednotky. Magnetickou pásku lze učinit trvale a bezpečně odolnou proti zápisu a přepisu pouze mechanickým přestavěním zámku na vlastní kazetě.

MAID

Technologie **MAID** (Massive Array of Inactive Disks) je zajímavá z hlediska využití levných pevných disků, které nejsou konstruovány pro trvalé zatížení v diskových polích. Diskové pole takovýchto disků je většinu provozní doby ve vypnutém stavu, hlavy disků jsou zaparkovány a diskové svazky se neotáčí. Teprve ve chvíli příchodu požadavku na přístup se disky roztočí a stanou se aktivními pro čtení a zápis. Je to jedna z variant dosažení velmi vysokých kapacit za nízkou cenu. V praxi zatím není dostatek informací pro stanovení spolehlivosti a trvanlivosti tohoto řešení.

Přechodové stavy, které vznikají zejména při aktivaci disků mohou mít negativní vliv na životnost mechanických částí.

Kazeta s pevným diskem

V pevné kazetě o tloušťce 1/2" je uložen pevný disk s charakteristickým rozměrem 2,5" a rozhraním **SATA**. Tato technologie by měla přinést zrychlení inicializace kazety v mechanice asi 10x při zachování možnosti vyjmout kazetu z knihovny a uložit na bezpečné místo. Použití 2,5" diskové technologie počítá s vysokou odolností těchto menších pevných disků proti otřesům. Jednotka pro čtení kazet má charakteristické rozměry 5,24" páskové jednotky a výrobce předpokládá, že bude snadné ve stávajících robotických knihovnách přejít na tuto technologii.

Virtuální knihovna

Hardwarové virtuální knihovny - přináší kombinaci dvou konvergujících světů - levného úložného prostoru na diskových polích s disky **SATA** a páskové robotické knihovny s výhodnou cenou za uložený **TB** (terabyte). Takováto virtuální knihovna se navenek do **SAN** prezentuje jako skutečná a její nasazení nevyžaduje ani minimální změnu v konfiguraci zálohovacího a archivačního prostředí. Uživatel získává rychlost diskového pole za cenu redukovanou použitím páskových médií. Interní přesuny dat mezi diskovým polem a páskovými médii jsou řízeny mechanismy **ILM** (Information Lifecycle Management).

Zálohování archivů

Problematika zálohování archivu narůstá především s objemem uložených dat. Stále více se do popředí dostávají technologie časových snímků metadat na levných diskových kapacitách. Klasický archiv na páskových médiích dokáže využít vlastností paralelního toku dat do několika zápisových jednotek a tím vytvářet bezpečnostní klony pásek přímo v době vzniku. Klonování může probíhat mimo hlavní čas využití zařízení. Tato technologie je u diskových polí nepoměrně dražší a složitější. Disková pole častěji využívají technik vzdálených zrcadlení nebo replikací. Tento způsob ochrany dat se rozšířením po teroristických útocích v září 2011 v USA stává levnějším a dostupnějším

a není vyloučeno, že do značné míry nahradí i klasické zálohování. Jistě to bude i dobré řešení pro dlouhodobé archivace elektronických dat.

Datové úložiště musí mít vypracováno plán obnovy po havárii. V případě ztráty technického prostředku v jedné lokalitě nastoupí náhradní datový zdroj v lokalitě druhé automaticky, aniž by musel být nutně dodržen parametr rychlosti odezvy u uživatelů. V případě ztráty celé lokality ji opět automaticky nahrazuje lokalita další a původní lokalita je rekonstruována podle schváleného plánu obnovy po havárii. Priority obnovy budou závislé na způsobu použití jednotlivých úložišť. Způsobem využití se myslí režim primární + záložní nebo paralelní užití všech najednou. V prvním případě je havárie záložního úložiště podružná, v druhém případě je havárie kteréhokoliv úložiště kritická, protože všechna úložiště jsou využívána paralelně z důvodu jejich maximálního vytížení.

2.6 Doporučený model řízení archivace na dobu stovek let

Hlavním cílem dlouhodobé archivace je prakticky nekonečně dlouho fungující úložiště dat s takovým stupněm zabezpečení, který zaručí jejich trvalou nezničitelnost.

Nejdůležitější zásady fungování jsou shrnuty v následujících bodech:

1) Stávající technologie a jejich dosavadní vývoj vyžadují trvalý proces migrace dat z jedné generace na generace následující.

- V budoucnu může být vyvinuta dosud neznámá technologie se schopnostmi dlouhodobě skladovat informace bez zvláštních nároků na metody přístupu k informacím.

2) Na dlouhodobý archiv nelze aplikovat mechanismy **ILM**

- Data v archivu se vyznačují tím, že vysoké procento informací nebude z archivu již nikdy požadováno.

3) Zařízení a média pro úschovu dat musí mít výrobcem deklarovanou životnost minimálně 10 let

- Projektovaná životnost **IT** (Information Technology) je zpravidla pouze pět let. Deset let je doba morálního zastarání, za kterou již významně rostou ekonomické náklady na jejich provoz.

4) Data musí být uložena v nejméně třech lokalitách

- Třetí lokalita by měla fungovat jako trvalé úložiště typu **WORM**. Tato lokalita v žádném případě neslouží k primárnímu zpracování dat, pouze a jen k jejich trvalé úschově.

5) Alespoň v jedné lokalitě musí být použita zařízení z různé technologie, jejíž perioda obměny se nekryje s periodou obměny lokalit ostatních

- Výměna technologie nesmí ohrozit nepřetržitost existence dat.

6) Dvě lokality jsou synchronizovány v reálném čase, třetí lokalita v asynchronním režimu se zpožděním

- Zpoždění zabraňuje vědomému zničení dat ve všech lokalitách.

7) Uložená data mohou být chráněna proti zneužití nejen v úložištích, ale i v komunikačních kanálech

- Bezpečnost nemusí být vyžadována vzhledem k obsahu informací, ale s ohledem na nedovolené komerční využití.

8) Archiv nevyžaduje nasazení prostředků pro úschovu dat s vysokou dostupností

- Vysoká dostupnost prostředků několikanásobně zvyšuje jednotkovou cenu za uložený **TB**. Archiv není kritická aplikace. Kritická je dlouhodobá trvanlivost informací.

9) Zařízení pro úschovu dat a komunikace v jednotlivých lokalitách by neměla pocházet od jednoho originálního výrobce **OEM** (Original Equipment Manufacturer)

- Vazba na jednoho výrobce je riskantní. Neustále probíhající konsolidace trhu **IT** způsobuje pohlcování jednotlivých článků světové produkce a často to znamená i ukončení podpory.

10) Zařízení pro úschovu dat musí být co nejvíce otevřené obecně rozšířeným platformám bez úzké vazby na jednoho dodavatele

- Aplikační část archivu musí být zcela nezávislá na hardwarové platformě.

11) Je nutno vyhnout se unikátním technologiím, za kterými stojí pouze jeden výrobce, i s rizikem, že vybrané řešení nebude nejvýkonnější

- Stále platí zkušenost z minulosti, kdy se hledal univerzální a široce akceptovaný formát videozáznamu mezi principy **BETA** (je záznamový standard firmy SONY pro televizní obraz a zvuk ke komerčnímu využití) a **VHS** (je záznamový standard firmy JVC pro televizní obraz a zvuk komerčního využití). Úspěšnější byl nakonec méně dokonalý **VHS**.

12) Prvotní pořizovací náklady a cena za implementaci bude tvořit pouze desetinu až třetinu celkových nákladů na provoz a vlastnictví.

- Z ekonomického hlediska se pro dlouhodobý projekt jeví nejvýhodnější plný outsourcing, pokud není na závalu dislokace dat mimo území jejich správce. Světové trendy však jednoznačně směřují k tomuto modelu a s ním roste i propracovaná legislativa, ošetřující otázku vlastnictví a zodpovědnosti poskytovatele služby.

V současnosti se jeví jako výhodné použít jako úložiště diskové pole, založené na technologii **SATA**. Při dodržení výše uvedených zásad budou pořizovací náklady tak nízké, že jejich budoucí obnova nebude mít zásadní vliv na jednotkovou cenu za **TB**. Snížení nákladů na kapacitu lze dále dosáhnout kombinací zařízení založených na pevných discích a na výměnných páskových médiích. Případně lze využít hybridního systému, který samostatně řeší vnitřní organizační strukturu dat a navenek funguje jako automatizovaná knihovna s velkou kapacitou. Kromě výhodné jednotkové ceny za **TB** jde o kombinaci dvou technologií s různou dobou morálního zastarávání a při inovaci

jedné části budou data uložena v druhé části nebo na výměnných médiích. Budoucí vývoj naznačuje, že opět dojde k rozevření průběhu křivek cen za uloženou jednotkovou kapacitu na pevném disku a na páskovém či jiném výměnném médiu. To by opět znamenalo vytváření levných velkokapacitních úložišť, založených na výměnných technologiích.

Klíčovým momentem pro trvale udržitelný chod velkokapacitního distribuovaného archivu je rychlá a spolehlivá komunikace mezi pracovišti, bez ohledu na to, jaké prostředky jsou použity pro ukládání elektronických informací.

2.7 Technologie RAID obecně

Vysoký výkon a vysoká kapacita **RAID** (Redundant Array of Inexpensive/Independent Disks) je v informatice metoda zabezpečení dat proti selhání pevného disku. Zabezpečení je realizováno specifickým ukládáním dat na více nezávislých disků, kdy jsou uložená data zachována i při selhání některého z nich. Úroveň zabezpečení se liší podle zvoleného typu **RAID**).

RAID 0 – má funkci, která rozděluje data do bloků a každý z bloků je zapsán na zvláštním disku. Toto uspořádání pracuje od dvou jednotek (disků). Jednoduše řečeno data plynou přes mnoho kanálů na různé jednotky. Dosažení nejlepšího výkonu je možno za předpokladu, že každá jednotka má svůj řadič. Toto nastíněné řešení má největší přenosovou rychlost ze všech polí.

RAID 1 – zapisuje data vždy na dva disky a vzájemně se zrcadlí. Daný řadič musí provádět simultánní zápis na zrcadlené jednotky a zároveň musí umět paralelní čtení z jednotek. Tudíž doba zápisu je stejná jako na jednu jednotku. Potom doba čtení je poloviční a přístupová doba stejná. Tímto řešením dochází k totální redundanci dat. Při možné chybě nebo poškození jednotky jsou uložená data nehrožena a poškozená jednotka se jednoduše vymění s následným kopírováním chybějících dat. Jeden datový kanál pro dva disky se nazývá mirroring a použití dvou datových kanálů se jmenuje duplexing.

RAID 2 – je založen na základu **RAID 0**. Do pole se přidávají diskové jednotky, které mají v sobě zabudovanou dodatečnou ochranu dat pomocí **ECC** korekce (Error Checking and Connection). Zmíněná korekce vyžaduje pomoc ze strany pevných disků. Kontrolní informace jsou vytvářeny při počátečním zápisu a při čtení se pak výstup podle ní průběžně kontroluje, čímž jsou poli umožněny vysoké přenosové rychlosti. Čím více rychlost roste, tím je nutných více datových disků, ale klesá poměr datových/kontrolních disků a tím pádem jsou snižovány náklady na pole.

RAID 3 - je stejně jako **RAID 2** založen na základu **RAID 0**. Na rozdíl od **RAID 2** se používá sofistikovanější a efektivnější metoda ochrany integrity dat. Paritní informace je ukládána na vyhrazený disk. Pro sestavení pole jsou nutné nejméně tři jednotky a data jsou na disk zapisována po několika bajtech. Když dojde k náhlému selhání jednotky a její následné nutné výměně, tak existuje možnost ztracená data dopočítat z paritních informací. Naproti tomu vytvářet paritní informace snižuje rychlost zápisu. Výkon při čtení v případě **RAID 3** je patřičně vyšší než při čtení z jednoho disku. Při čtení malého objemu dat je rychlost čtení **RAID 3** blízká rychlosti čtení z jedné jednotky.

RAID 4 – je přebraný od **RAID 3**. Hlavní rozdíl je ve změně výpočtu paritní informace. Naproti počítání parity z bitů je v tomto případě parita výsledkem paritní funkce nad bloky dat. Tím pádem je obrácena výhodnost mezi čtením velkých a malých bloků dat. Je efektivnější čtení velkého množství malých bloků, ale rychlost zápisu je nízká.

RAID 5 – datové diskové jednotky s distribuovanými bloky paritních informací, kde k sestavení pole jsou zapotřebí nejméně tři jednotky. Překonává nedostatky **RAID 3** a **RAID 4**. Paritní informace se neukládá na vyhrazený disk, ale je rozložena na všechny jednotky pole. Rychlost zápisu je rychlejší než u **RAID 3** či **RAID 4**. Při zápisu se musí neustále číst paritní informace a přepočítané je následně ukládat. Zápis je v **RAID 5** pomalejší než v **RAID 0**. Výkon při čtení se modifikuje nastavením velikosti bloku ukládaných dat pro aplikaci, která je používána. Rozdíl ve výkonu při čtení a zápisu je řešen rozšířením pole metodami opožděného zápisu a implementací v systémech s více procesory.

RAID 6 – datové diskové jednotky se zdvojeným distribuovaným paritním blokem. Jedná se o rozšíření konceptu **RAID 5**. Pole má vyšší odolnost vůči chybám, protože je použito druhé ukládání, nezávislé paritní informace. **RAID 6** je proto nejspolehlivější a při výpadku obou disků je možné data znovu nahrát. Rychlost čtení se dá srovnat s **RAID 5**, ale zápis je pomalejší, protože se provádí výpočty a ukládají se dvě sady paritních informací. Cena je též vyšší, proto se používá jen v případě maximální nutnosti, aby data byla velmi spolehlivá a rychle přístupná.

RAID 7 - jedná se o soukromě vytvořený patent společnosti Storage Computers. Funguje jako asynchronní systém s vlastní vyrovnávací pamětí a nezávislým řízením každého disku. Pole funguje podobně jako osobní počítač s mikroprocesorem a vlastním operačním systémem. Všechny vstupní/výstupní informace jsou nezávisle kontrolovány a ukládány do vyrovnávací paměti. Všechna čtení a zápisy jsou navíc ukládány do vyrovnávací paměti centrálně. Soustava obsahuje minimálně jeden paritní disk. Na systém lze napojit 48 disků a 12 hostitelských počítačů. **RAID 7** je schopen odhalit i sudý počet chyb, které si u předešlých systémů zachovávají paritu a není možné je objevit. Generování paritních informací je integrováno s vyrovnávací pamětí. Pár jednotek může být v systému označeno za nepřipravené.

Avšak je nutné si uvědomit, že **RAID** nenahrazuje zálohování dat, kompatibilitu s prostředím **Apple** (americká, nejen počítačová, firma) a produkty **VMWare**, které slouží pro virtualizaci jednoho nebo i více počítačů na jednom hostitelském stroji. **VMWare** je členem Linux Foundation.

2.8 Cloud computing

Cloud computing je na internetu založený model vývoje a používání počítačových technologií. Lze ho také charakterizovat jako poskytování služeb či programů uložených na serverech na Internetu s tím, že uživatelé k nim mohou přistupovat například pomocí webového prohlížeče nebo klienta dané aplikace a používat prakticky odkudkoliv. Uživatelé neplatí (za předpokladu, že je služba placená) za vlastní software, ale za jeho užití. Nabídka aplikací se pohybuje od kancelářských

aplikací, přes systémy pro distribuované výpočty, až po operační systémy provozované v prohlížečích, jako je například **eyeOS, Cloud** či **iCloud**.¹

Cloud computing je z principu jakýkoliv program nebo služba, která je dostupná z jakéhokoli místa, z domácího i pracovního počítače, nebo z internetové kavárny, protože ukládá uživatelská data nebo celou svou aplikační logiku na server kdesi na internetu. Co z této definice plyne? Cloud computingem je i běžný e-mail ale ostatně i webová stránka s osobním profilem, který je dostupný z jakéhokoliv počítače na celém světě.²

2.9 Amazon Web Services

AWS jsou různé webové služby společnosti Amazon založené na technologii Cloud Computing. Nejznámějšími službami jsou pravděpodobně Amazon EC2 a Amazon S3. Amazon EC2 půjčuje klientům výpočetní výkon Cloudu, aby jej mohli použít na své vlastní aplikace. Amazonu se hradí částky za využitý výkon a za přenos souborů. Amazon S3 funguje podobně, ale namísto výpočetního času propůjčuje klientům úložiště. Mezi nejznámější klienty patří různé webové aplikace pro sdílení obrázků či služby Twitter nebo SlideShare.³

Na našem území s podobnými službami začíná například Telefonica O2, ovšem do propracovanosti Amazonu je velmi vzdálena.⁴

¹ *Cloud Computing* [online]. 2011. 2010 [cit. 2011-05-30]. Dostupné z WWW: <<http://www.cloud-computing.cz/>>.

² *Zive.cz* [online]. 2011. 2008 [cit. 2011-05-30]. Cloud Computing. Dostupné z WWW: <<http://www.zive.cz/Clanky/Cloud-computing-slibna-budoucnost-nebo-marketing/sc-3-a-144443/default.aspx>>.

³ *Amazon* [online]. 2011. 2011 [cit. 2011-05-30]. Ec2. Dostupné z WWW: <<http://aws.amazon.com/ec2/>>.

⁴ *O2* [online]. 2011. 2011 [cit. 2011-05-30]. Managed Backup and Restore. Dostupné z WWW: <http://www.o2.cz/corporate/198026-managed_services/114812-managed_backup_restore.html>.

3 ANALÝZA PROBLÉMU A SOUČASNÁ SITUACE

3.1 Vybraný případ pro provedení zálohování dat

Zákazník - výrobní firma střední velikosti (komponenty osobních automobilů - 250 zaměstnanců)

Lokalita - Jihomoravský kraj s centrálou v Brně a 5 pobočnými závody v okolí do 70km

Operační platforma - MS Windows a Linux

Objem dat - 4TB celkem

Objem denních změn - 600GB

Požadavky na retenci (zachování) - 12 měsíců

Charakter provozu IT - pracovní týden administrativa v jedné směně, výroba nepřetržitě ve dvou směnách

3.2 Současný stav

Firma využívá následující prostředky IT:

- *server elektronické pošty* – slouží pro vzájemnou komunikaci uživatelů pracujících ve firmě pomocí počítačů, souborů, internetu a dalších komunikačních linek. E-pošta je označení pro komunikaci v rámci internetu či intranetu. Když chce uživatel pracovat na internetu, tak musí mít k dispozici poštovní program podporující protokol (**POP3**, lépe **IMAP4**) pro přijímání pošty a (**SMTP**) pro odesílání pošty. Dále musí mít zřízenou poštovní schránku u daného poskytovatele internetového připojení (**ISP**), nebo u internetového poskytovatele, který nabízí freemailové nebo placené služby. Každý uživatel má svoji e-mail adresu, uživatelské jméno (obvykle část e-mail adresy) a heslo pro odchozí poštu. Firma má vlastní server MS- Exchange 2007 v uspořádání

cluster, který je základem portfolia Microsoftu v oblasti nabídky firemních systémů. Počítačový cluster je seskupení počítačů, které mezi sebou mohou velice úzce spolupracovat, což znamená, že navenek se jeví jako by pracoval jen jeden počítač. Musí být samozřejmě propojeny počítačovou sítí.

- *databázový SQL server* – firma používá databázový systém od už několikrát zmiňované a proslulé společnosti Microsoft, který se nazývá Microsoft SQL Server, jež slouží pro informační systém řízení výroby a pro **CRM** (Customer Relationship Management – databázová technologie podporující procesy shromažďování, zpracovávání a následného využívání informací ohledně zákazníků firmy). Microsoft SQL Server má firmě pomoci snižovat celkové náklady a zvyšovat efektivitu organizace práce, protože jasněji a přehledněji zpracovaná data ulehčí jak práci pracovníkům, tak komunikaci s vrcholným vedením a následně s dodavateli či odběrateli.
- *databázový INFORMIX server* – neboli Informix Dynamic Server (**IDS**) je dosti výjimečný databázový produkt americké společnosti v oboru informačních technologií (**IBM**). Mezi jeho největší přednosti patří spolehlivost, dobrá ovladatelnost a pečlivé zpracování datových převodů. **IDS** se dá použít pro různé druhy aplikace. Firma jej využívá pro ekonomické řízení, kde je důležité, že nezáleží na objemu dat, protože s každým rokem se objem dat zvětšuje. **IDS** je snadno implementován na všechny běžné operační systémy a je též dostupný v několika verzích.
- *zálohovací server s robotickou páskovou knihovnou* – firma používá robotickou knihovnu **ADIC Scalar 100 LTO**, která je vysoce výkonná a škálovatelná. Tato robotická knihovna obsahuje pásky v měřítku od 15 až po 72 a jeden až šest disků v závislosti na výběru technologie a konfigurace. **ADIC Scalar 100** podporuje **LTO-1, LTO-2, AIT-2, AIT-3, SDLT220, SDLT220, SDLT320** pásky, ale firma využívá **LTO-1** pásky. Přidávání kapacity je snadné. Upgrade může trvat méně než 10 minut bez žádné změny regálového systému nebo přidáním výtahů či pass-through portů vevnitř zařízení na lepší manipulaci mezi kazetami. Tento model nabízí 18 slotů dat a 1 slot poštovní schránky. Za zmínku stojí nezanedbatelné vlastnosti zjednodušující případné budoucí změny, jako

jsou: nastavitelná robotika pro všechny modely, schopnost skladování si vzájemně poskytovaných služeb a efektivní spolupráce se systémem tzv. interoperabilitu, přidávání nových pásků podle potřeby, diagnostické možnosti jako jsou kontroly stavu systému v reálném čase, email nebo pager upozornění a telefonických výpisů až do domu. Využívá technologie Backup Tape Drives, což je záložní systém. Páskové knihovny se v současnosti připojují buďto skrze **SCSI to FC** nebo samotné mechaniky disponují konektivitou native fiber channel. Ve firmě se používá prvně zmíněný způsob skrz standardní rozhraní (**SCSI** – Small Computer System Interface) pro výměnu dat mezi externími či interními počítačovými zařízeními a počítačovou sběrnici.

Obrázek 1: Robotická knihovna ADIC SCALAR 1000 LTO (Zdroj: autor a management společnosti)



Obrázek 2: Robotická knihovna ADIC SCALAR 100 LTO (Zdroj: autor a management společnosti)



- *server s technologií VoIP-* (**VoIP-** Voice over Internet Protocol) umožňuje přenos digitalizovaného hlasu v těle paketů tří druhů protokolů **UDP/TCP/IP**, patřící mezi základní protokoly na internetu, prostřednictvím počítačové sítě či nějakého jiného média, postupného pro protokol **IP**. Hlavní využití je pro autonomní hlasovou potřebu prostřednictvím internetu, intranetu nebo jakéhokoli jiného datového spojení, čehož firma bohatě využívá. Jedna z nutných podmínek pro srozumitelné a spolehlivé telefonní spojení přes technologii **VoIP** je zajištění kvality služby, které se označuje (**QoS** – Quality of Service). V případě výskytu notebooku je nasazeno bezdrátové *VoWLAN* (*Voice over Wireless LAN*). Firma využívá počítačového softwaru s otevřeným zdrojovým kódem, který nese název **ASTERISK**. Asterisk je aplikace, která implementuje telefonní ústřednu pomocí obyčejného počítače. Mezi jeho největší výhody patří ušetření kabeláže, lepší nastavitelná konfigurace bez potřeby technika. Jedná se o jedno z nejrozšířenějších řešení v oblasti integrovaného telekomunikačního softwaru. Je nabízen zejména telekomunikačním operátorem Telefónica O2 a běží na platformě Linux. Firma se rozhodla pro jiného zprostředkovatele, a tím je společnost Faster CZ, protože má v současnosti nejkvalitnější servis, jež je šitý na míru a nastavitelný

požadavkům firmy s mnoha pobočkami po celé republice, kde je vyžadována záloha proběhlé komunikace.

- *server pro přístupový a docházkový systém* – firma používá samostatný server pro sledování docházky přes zdarma dostupný databázový systém Microsoft SQL Express Edition. Jedná se o odlehčenou verzi populárního Microsoft SQL Server s určitými limity, na něž budou narážet jen opravdu nároční uživatelé a pro potřeby docházek a přístupu je více než dostatečná. Pro případ přechodu na placenou verzi se firma nemusí obávat žádných zásahů do již vytvořeného systému.
- *centrální datové úložiště se souborovým přístupem* – je používán, klasický File Server, kde se ukládají dokumenty volně ke sdílení mezi registrovanými pracovníky neboli uživateli systému. Mimo jiné se zde nachází naskenovaná dokumentace, firemní doklady, faktury, fotky a jiné. Časem nepotřebná dokumentace je skladována externě na fyzickém úložišti u firmy Royal Seal pana inženýra Beránka.
- *souborové servery na pobočných závodech* – fungují na stejném principu jako na centrálním datovém úložišti. Na každém z nich běží account system, který v podstatě určuje přístupová práva přidělená správcem systému. Zjednodušeně, každý pracovník má přístup jen k určité aplikaci systému. Dále je zdůrazněno rozdělení pobočky dle funkce, kterou jí přidělí centrála, např. logistika, výrobní závod, export, import, atd.
- *komunikační servery na pobočných závodech* – zajišťují kryptovanou komunikaci mezi pobočkami, aby nebyla veřejně dostupná konkurenčnímu okolí. Jsou založeny na principu ochranného systému Firewall.
- *pracovní stanice a přenosné počítače jednotlivých pracovníků* – vybavení pracovních stanic je úměrné postavení zaměstnance. V případě vyšší funkce se mu do ruky dostává notebook, který může využívat i mimo pracovní dobu. Každý notebook je vybaven speciálním systémem na zálohování dat, neboť v minulosti docházelo k častým ztrátám citlivých informací.

- *IT infrastruktura hlavního závodu* - sestává ze strukturované kabeláže s aktivními prvky 100Mb/1Gb. 100Mb je rozveden po celém sídle firmy a páteřní síť pro komunikaci mezi servery má 1Gb. Středisko se nachází v centrální serverové místnosti se samostatnou datovou sítí **SAN**, s několika přepínači FC 2Gb a diskovým polem s rozhraním 2x FC 2Gb. Vybavení doplňují další disková pole s rozhraním iSCSI pro režijní data.
- *vybavení pobočných závodů* - se sestává ze strukturované kabeláže s aktivními prvky 100Mb a servery v přístrojové skříni, umístěné mimo výrobní a kancelářský provoz.
- *retence* - mnohé organizace narážejí na problém uchovávání dat po dlouhá časová období. Tato tzv. retenční data bývají uložena v archivech nebo skladech. Tato data musí být skladována z důvodu legislativních předpisů, a je nutno mít potřebné dokumenty a záznamy vždy k dispozici. Obecně správa a řízení spravovaných retenčních dat může představovat značný problém, pokud vezmeme v úvahu i další faktory, jako jsou náklady, integrace, výkon, zabezpečení, důvěrnost a období pro uchovávání na bázi zásad.

Zálohování provozních dat na centrále je prováděno pomocí zálohovacího serveru, s připojenou robotickou knihovnou. Zálohovací program provádí jedenkrát týdně plnou zálohu vybraných serverů a v ostatní dny ukládá pouze rozdílová data. Týdenní zálohy mají retenci jeden měsíc, měsíční zálohy vykonávané na konci posledního týdne v měsíci mají retenci jeden rok.

Data pobočných závodů jsou zálohována pomocí nativních nástrojů OS na jednotlivých serverech do vestavěné páskové jednotky DAT. Systém záloh je nastaven na denní plné zálohy s týdenní retencí (zachování).

3.3 Důvody nevyhovujícího systému

- *nedostatečná kapacita s ohledem na stále rychlejší růst objemu dat* – zavádění stále nových technologií generuje větší objem dat. Datová velikost roste trendově exponenciálně. Jsou zapotřebí neustále nové programy s nutností dívat se nepřetržitě do budoucna, které budou vyžadovat čím dál větší kapacity. Při zmínění nových technologií musíme počítat s koupí lepšího procesoru, serveru, paměti a síťových karet.
- *nedostatečný výkon zálohovacího serveru s robotickou knihovnou* – robotická knihovna je zastaralá a může pracovat maximálně s páskami LTO-2. Její kapacita je též nedostatečná, má velikost pouze sto pásek a nelze stavět robotické knihovny na sebe jako u jiných lepších verzí, tím pádem je potřeba zakoupit knihovnu s vyšší kapacitou, nebo se zaměřit na jinou technologii zálohy dat.
- *nesourodá média pro zálohování centrály a poboček* – poněvadž je centrála vybavena páskami LTO-2 a pobočky jsou vybaveny digitálními páskami DAT, proto se musí zálohovat zvlášť a je nemožná výměna dat mezi sebou, tudíž velmi těžko mezi sebou komunikují.
- *nedostatečná ochrana dat v jednotlivých lokalitách proti živlu a pracná obnova provozních dat po havárii* – nutnost uložení dat na dvou rozdílných místech pomocí replikací, například za pomoci VMware. K tomuto uvědomění došlo zejména po velkých povodních na konci devadesátých let a na počátku 21. století. Dříve se pásky s daty odvážely do bank, kde se s nimi zacházelo jako s ceninami. V dnešní době je kladen důraz na kryptování veškerých dat na nosičích, které se mohou nacházet mimo vlastní organizaci. Patří k nim data na páskách, flash discích a na pevných discích notebooků. V případě obnovy po havárii nastává složitý proces rekonstrukce kryptovaných provozních dat, a to kvůli možné ztrátě klíče. Klíče je tedy také chránit proti znehodnocení, ztrátě či zneužití. Každodenní zálohy by se měly vytvářet zejména pro obnovu stavu po katastrofické ztrátě dat a informací, tzv. Disaster Recovery s Boot médiem, které se po havárii původního OS spustí jako při nové instalaci (nejčastěji DVD). V

procesu obnovy se tedy nejprve obnoví OS, poté instalované programy, včetně programu pro zálohování a nakonec zálohovaná data. Firemní Disaster Recovery Plan je nutné podrobovat pravidelnému ověřování (pololetně nebo kvartálně), zda plánovaná obnova probíhá podle předpokladů. Cílem daného testu je posouzení míry splnitelnosti aktuálních firemních požadavků na obnovu po havárii. S každým novým komponentem či hardwarem je nutné vytvořit Boot medium a aktualizovat Disaster Recovery Plan.

- *nevyhovující* **RPO** (Recovery Point Objective) - popisuje přijatelnou ztrátu dat v měřeném časovém úseku. Jedná se o časový bod (před havárií) stanovený individuálně dle každé společnosti, ke kterému se obnoví data. Vystihuje je otázka: Kolik dat si můžeme dovolit ztratit v případě výpadku či havárie? Obecná definice říká: maximální požadovaný interval před výpadkem či havárií, během kterého lze v důsledku obnovení přijít o časové změny dat. Změny dat předcházejí výpadek nebo havárii před minimálně tímto intervalem jsou zachovány obnovením. Nula je platná hodnota a odpovídá požadavku na „nulovou ztrátu dat“. **RPO** dovoluje společnosti nastavit si časový prostor před možnou havárií, při níž dochází ke ztrátě dat. Hodnota údajů v tomto časovém prostoru je poměřována oproti preventivním nákladům, které by bylo nutné vynaložit v daném časovém úseku. **RPO** je nezávislý na čase, který je potřebný k uvedení nefungujícího systému zpět do stavu on-line. Pokud RPO společnosti je dvě hodiny, pak při nastartování systému po havárii musí být všechna data uvedena do původního stavu, i když společnost bere na vědomí, že data z časového úseku dvou hodin před katastrofou mohou být ztracena, tudíž přijatelná ztráta je dvě hodiny.
- *nevyhovující* **RTO** (Recovery Time Objective) – Jak dlouho můžeme postrádat jednu či více aplikací či dokonce systém? Tuto dobu lze definovat jako maximální požadovanou dobu nutnou k obnovení jedné či více aplikací a s nimi spojených dat do řádně funkčního stavu. Jinými slovy se jedná o cílový čas obnovy, v níž musí být podnikové procesy obnoveny po havárii (nebo narušení), aby se zabránilo nepřijatelným důsledkům spojených s přerušením kontinuity podnikání. Zahrnuje to čas pro pokusy vyřešit problém bez obnovy, samotnou

obnovu, testy a komunikaci s uživateli. Časová osa kontinuity podnikání často běží paralelně s časovou osou správy nehod. **RTO** je stanoven během analýzy podnikatelského dopadu **BIA** (Business Impact Analysis) vlastníkem procesu. **RTO** je potom prezentováno vrcholovému vedení k přijetí. **RTO** se váže k podnikovým procesům a nikoliv ke zdrojům potřebným k podpoře procesu. **RTO** a výsledky **BIA** slouží v plném rozsahu jako základ pro identifikaci a analýzu strategie uskutečnitelnosti začlenění do plánu kontinuity podnikání. Strategii uskutečnitelnosti by patřily jakékoliv možnosti, které by umožňovaly obnovení podnikových procesů v časovém rámci blížícím se **RTO**. To bude zahrnovat alternativní řešení nebo manuál s postupy, které nutně vyžadují počítačové systémy ke splnění **RTO**.

4 VLASTNÍ NÁVRHY ŘEŠENÍ, PŘÍNOS NÁVRHU ŘEŠENÍ

4.1 Možnosti řešení

Hitachi data systems (HDS)

Distribuce produktů světového výrobce zařízení pro ukládání dat. Škálovatelnost paměťových systémů dle potřeb zákazníka dohromady s centrálním ovládáním všech **HDS** zařízení v síti prostřednictvím **HITACHI** (japonská firma specializovaná na vyspělou a náročnou technologii) software, staví tyto produkty na špičku dnešní světové nabídky.

SGI

Distribuce výkonných serverů na Intel platformě a storage zařízení americké společnosti Silicon Graphics Inc. **SGI** dodává kompletní nabídku vysoce výkonných serverových a paměťových řešení spolu s pokrokovými profesionálními službami a podporou, které umožňují zákazníkům překonat výzvy komplexních intenzivních datových procesů a zrychlit objevy, inovace a informační přeměny.

INFORTREND

Distribuce diskových polí. Infortrend nabízí ucelené portfolio diskových polí modelů **FC-FC** (Fibre Channel-**FC** je až gigabitové komunikační rozhraní, primárně používané pro vysokorychlostní přenos dat v síti. **FC** se začalo používat především v oblasti superpočítačů, ale stává se též standardním typem připojení na sítích **SAN**. **FC** může být použito i na kroucené dvojlince měděného drátu nebo na optických kabelech).

FC-SAS (Serial Attached SCSI - **SAS** je v informatice sériová sběrnice, která nahrazuje paralelní **SCSI** sběrnici. Slouží k připojení pevných disků a páskových jednotek).

SAS-SAS, iSCSI-SAS (Internet Small Computer System Interface - **iSCSI** je v informatice síťový protokol, který umožňuje připojovat úložný prostor, což jsou nejčastěji disková pole, pomocí počítačové sítě).

FC-SATA (Serial ATA -**SATA** označuje v informatice počítačovou sběrnici, která využívá datové rozhraní pro připojení velkokapacitních paměťových zařízení, jako jsou pevné a optické disky. V dnešní době je **SATA** řadičem vybavena většina stolních a přenosných počítačů).

SCSI-SATA (Small Computer System Interface - **SCSI** je standardní rozhraní a sada příkazů pro výměnu dat mezi externími nebo interními počítačovými zařízeními a počítačovou sběrnici. Obvykle se používá pro připojení pevných disků nebo magnetopáskových jednotek. Lze také připojit i jiná zařízení např. skenery, jednotky **CD-ROM** nebo **DVD**).

Dále také varianty s duálními kontrolery a expanzními boxy **JBOD** (Just a Bunch of Disks - jsou jednotlivá zařízení, která jsou spojena do logického celku a vytváří tak kapacitu součtu všech členů).

Disková pole jsou standardně vybavena funkcí **„snapshot“** (označuje v informatice stav systému v určitém časovém okamžiku. Analogická funkce existuje ve fotografování. Může tedy odkazovat na aktuální kopii stavu systému nebo na určitou schopnost poskytnutou daným systémem).

ISILON

Distribuce unikátní clusterové technologie **NAS** (Network Attached Storage - je datové úložiště připojené k místní síti **LAN** - local area network. **NAS** úložištěm může být server nebo tomu věnované zařízení, které obsahuje jeden nebo více disků), jejímž jádrem je operační systém **OneFS** (je distribuovaný souborový systém navržený Isilon Systems). Používá se pro **Isilon Q** aplikace datových úložišť. Výjimečnost klastrů **Isilon IQ** spočívá ve skutečnosti, že s rostoucí kapacitou roste i výkon. Klastř může dosáhnout kapacity více než 1.6 PB s maximálním agregovaným výkonem 7Gb/s. Určeno především pro práci s multimediálními daty o velkém objemu.

PETASTOR

Distribuce diskových polí pro široké použití za cenově výhodných podmínek. Diskové systémy určeny pro **DAS** (Direct-attached storage - je digitální úložný systém přímo propojený k serveru nebo pracovní stanici bez datové sítě mezi nimi. Je to

retronym - druh neologismu, který zavádí nový pojem odlišný od původní verze , který se používá především k odlišení bezsít'ových úložišť od **SAN** a **NAS**.).

SAN prostředí (Storage area network - je datová síť oddělená od **LAN**, **WAN** atp., která slouží pro připojení externích zařízení k serverům). **SAN** vznikla hlavně kvůli narůstajícím potřebám na zabezpečení a konsolidaci dat. Díky poměrně vysokým pořizovacím nákladům se **SAN** budují hlavně ve větších společnostech. Vysoký výkon a odolnost celého systému vůči chybám (**RAID6**).

QUANTUM

Distribuce zálohovacích knihoven, autoloaderů a mechanik, včetně velkých profesionálních knihoven s mimořádnou škálovatelností, s kapacitou desítek **PB** (petabyte) a vestavěnou inteligencí. Nabídka zahrnuje i virtuální knihovny s deduplikací (komprese dat až do poměru 1:60).

TANDBERG DATA

Distribuce cenově výhodných **DLT** (Digital Linear Tape - je záznamové médium pro velkokapacitní zálohovací zařízení. Používá technologii magnetického pásku pro záznam a uchování počítačových dat).

SLR mechanika (Scalable Linear Recording - je v informatice technologie, která se používá pro záznam dat na magnetické pásky. Termín **SLR** je nyní často používán v souvislosti s **QIC** - báze půlpalcové kazety, jež jsou jediné vyráběné disky použitelné do zmíněné mechaniky).

LTO mechanika (Linear Tape-Open - je v informatice technologie, která se používá pro záznam dat na magnetické pásky. **LTO** je otevřený formát pro páskové produkty určené pro použití ve středně náročných systémech a na ostatních serverech).

Nabídka zahrnuje i malé automatizace počínaje jednou mechanikou a 8 sloty pro pásky až po knihovny vybavené až osmi mechanikami a 77 sloty a do kapacity 125 TB s kompresí.

FORTINET

Distribuce komplexních bezpečnostních systémů a systémů pro správu a reportování. V kategorii **SOHO** (small office/home office) až **Enterprise** řešení. Produkty firmy Fortinet nabízejí nejlepší volbu ve své třídě pro sedm klíčových bezpečnostních funkcí. Fortinet nabízí jednoduché licencování nezávislé na počtu uživatelů, náklady na následnou údržbu a aktualizace zařízení i technickou podporu jsou tak výrazně nižší než v případě konkurenčních řešení, která jsou licencována “na uživatele”.

STORNEXT

Distribuce software StorNext File System, který je navržen pro heterogenní prostředí pro sdílení systémů různých platforem na jednom diskovém prostoru. Sdílená data nemusí putovat skrze daný server a poté přes **LAN** na daný stroj, ale data lze sdílet přímo pro všechny servery bez ohledu na použitý operační systém.

NEXSAN

Distribuce diskových polí kategorie **SATA** pro využití v oblasti **Enterprise**, což je ukládání dat ve velkém měřítku. Pro srovnání se spotřebitelským úložištěm dat má vyšší škálovatelnost, vyšší spolehlivost, lepší odolnost proti chybám a mnohem vyšší počáteční cenu.

BABKONE

Distribuce zálohovacího softwaru založeného na modulární architektuře sloužící k archivaci a obnově dat. Je určen pro široké nasazení na platformách operačních systémů Unix a Windows, podpora síťových struktur **SAN**, **NAS**.

EMULEX, BROCADE, QLOGIC

Distribuce **FC** switch (síťový přepínač kompatibilní s **FC** protokolem., jenž umožňuje vytvářet **FC** konstrukce, které jsou v současné době hlavní součástí většiny **SAN**). Distribuce příslušenství.

VERBATIM

Distribuce vysokokapacitních paměťových médií, paměťových karet/čteček, flash disků a **USB** disků (Universal Serial Bus - je univerzální sériová sběrnice. Moderní způsob připojení periferií k počítači), **CD** (Compact Disc - je optický disk určený pro ukládání digitálních dat), **DVD** (Digital Versatile Disc nebo Digital Video Disc - je druh digitálního optického datového nosiče s výrazně větší kapacitou záznamu) a příslušenství.

czBackup

Prodej a pronájem systému automatizované zálohy využívající veřejné datové úložiště. Profesionální šifrování choulostivých dat, uživatelská přívětivost služby a její automatizovaný chod snižují nároky na odbornou obsluhu a tím i cenu zálohování. Zároveň jsou minimalizována rizika ztrát způsobených neodbornou manipulací.

4.2 Vlastní zálohování ve firmě

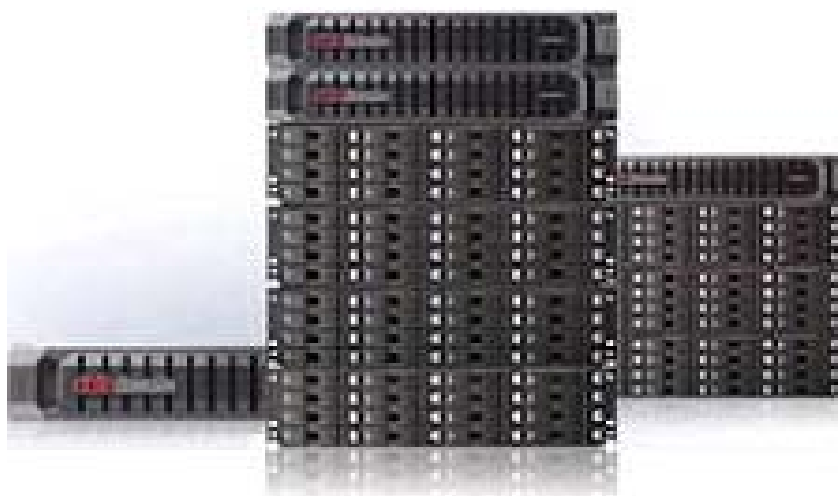
Virtuální knihovna

Nabízí se technologie autonomních virtuálních knihoven s možností replikace do vzdálených míst. Zachová se starý zálohovací systém se stejným programovým vybavením, ale na místo robotické knihovny se implementuje virtuální knihovna. Virtuální knihovna funguje prakticky stejně, jako fyzická robotická, je ale rychlejší, snadno se dá navyšovat její výkon a kapacita a nevyžaduje změnu zálohovacího programu ani strategie zálohování. Schopnost replikovat virtuální knihovny ve vzdáleném umístění řeší distribuci záloh do jiné lokality a tím i jejich ochranu proti živelným událostem. Nasazení virtuálních knihoven šetří také náklady na školení správců, kteří si nemusí osvojovat nové prostředí zálohování, ale použijí pro ně známé a ověřené prostředky původního programu pro zálohování.

Obrázek 3: Série zařízení Quantum Storage DXi s mimo jiné Quantum Dxi6520 (Zdroj: www.peripheralstorage.com/deduplication/quantum/quantum-dxi.htm)

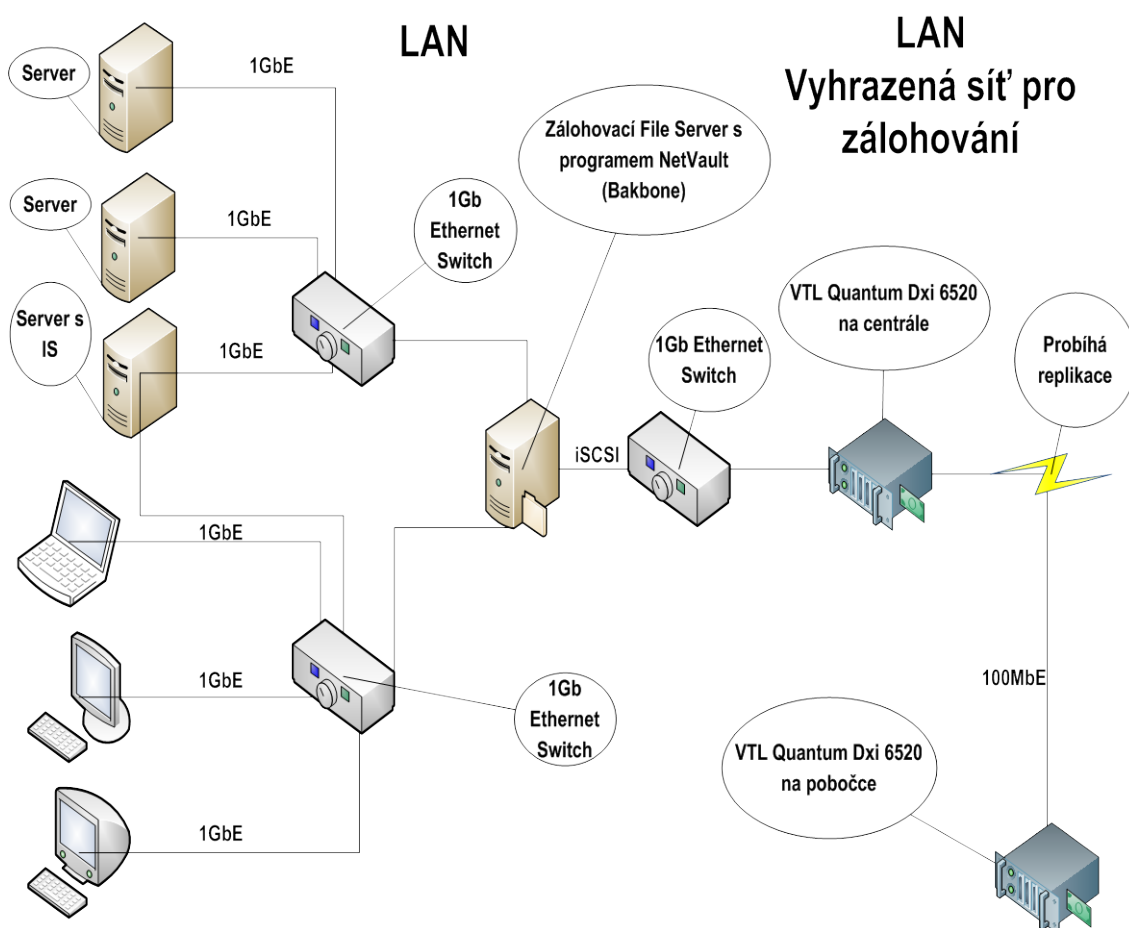


Obrázek 4: Série zařízení poslední generace od EMC Data Domain (Zdroj: www.altimate.pt/local/1b57de9018c00c17/Newsletter.html)



Modelový typ Quantum DXi6520 je dodáván v rozsahu kapacit 8 - 32TB, s maximálně šesti 1GbE porty. Nejlepší zvolenou konfigurací se jeví být s kapacitou 16TB, kterou lze do budoucna rozšiřovat po 8TB až na 32TB, poněvadž se musí počítat s vyhověním potřeb růstu firmy a s tím spojeným novým přívalem dat určených k záloze. S tak velkou kapacitou se najde rezerva i pro pobočky firmy. Nemí nutné se zbavovat staré robotické knihovny, neboť ji lze využít jako archivu, který zůstane zakonzervovaný z důvodů dlouholeté retenci dat.

Obrázek 5: Grafické schéma řešení virtuální knihovny (VTL) (Zdroj: autor a management společnosti)



Výhody:

Úspora nákladů na IT specialisty

Nově navrhovaný systém zálohování sjednocuje ukládání kritických dat do jedné platformy, dovoluje správu záloh z jedné aplikace a z jednoho místa. Odpadá tím organizační zabezpečení zálohování na pobočkách, kde často docházelo vlivem lidské chyby k opomenutí výměny média v mechanice a tím nabourání pravidelnosti a snížení spolehlivosti ochrany dat.

Využitím virtuálních knihoven se šetří náklady na školení **IT** pracovníků, kteří budou dále používat známé prostředí programu pro zálohování, změní se pouze fyzická podstata systému. Při využití externího datového centra pro distribuci záloh mimo mateřskou firmu se ušetří náklady na úschovu datových medií v odloučeném chráněném místě a odpadne organizace těchto medií.

Zvýšení bezpečnosti a spolehlivosti

Jednotná správa zálohování dovoluje lepší přehled o prováděných činnostech a rychlejší orientaci při vyhledávání požadovaných informací pro rekonstrukci. Přenos dat v kryptované podobě do externího datového centra vylučuje zneužití, ke kterému by mohlo docházet při transportu medií do vzdálené lokality.

Přechod na diskově orientované úložiště primárních záloh **VTL** zvyšuje faktickou spolehlivost celého procesu, využívající původně páskové technologie, které jsou mechanicky řádově složitější, než pevné disky. Samotná pásková média jsou velmi citlivá na zacházení, vadí jim prach, zvýšená teplota a mechanické nárazy.

Úspora energie

Moderní zařízení pro úschovu dat jsou energeticky méně náročná. Při nasazení **VTL** s technologií **MAID** (Massive Array of Idle/Inactive Disks) a **SSD** (Solid State Disk) dochází k výrazné úspoře elektrické energie, protože pevné disky v diskovém poli se v době nečinnosti postupně uvádí do režimu spánku (sleep) a roztáčí se znovu při dalším zálohování. V oblasti optimalizace správy dat se uspoří nejvíce energie a většinou zdarma pomocí smazání nepoužívaných dat a **ILM** (Information Lifecycle

Management), což je ukládání dat podle potřeb výkonu a zabezpečení. V oblasti optimalizace infrastruktury, kde je klíčem konsolidace. Konsolidují se přímo připojená úložiště do jednoho externího pole. V oblasti optimalizace zařízení se dají ušetřit finance použitím špičkových technologií.

Ochrana proti živlu i na pobočkách a vyřešení nesourodosti médií

Na pobočkách se použije technika replikací s centrálou. Při nynější situaci, že firma má několik poboček, se v každé z nich implementuje **VTL**. V nich probíhá obousměrná replikace, mezi centrálou a pobočkami. Nesourodost médií se automaticky vyřeší. Když se likvidují datová média **LTO**, je nutné udělat autorizovanou likvidaci. Autorizovaná likvidace je potvrzení o nemožné neužitelnosti dat konkurencí

Shrnutí výhod

Jak je napsáno výše, výhod **VTL** je mnoho. Těmi největšími jsou ušetření místa, energie, lepší chladicí systémy. Je dobré zmínit používání stejného zálohovacího programu, čímž odpadá nutnost zaškolení. To se bude týkat jen nové technologie **VTL**. Používání **de-duplikace**, což je zvláštní technika komprese dat, jež dokáže zabránit výskytu stejného bloku dat na jednom úložišti. Příkladem může být, **OS Windows 7**, kdy z každého počítače není nutné mít uložená data ve **VLT**, ale jen souhrn rozdílů v datech. Velikost redukce je mnohdy v desítkách procent. Zálohování trvá jen několik desítek minut.

Nevýhody:

- Cenová nevýhodnost v důsledku velké prvotní investice
- Vyšší nároky na internetovou propustnost dat
- Nutnost zakoupení optického propojení v důsledku velkého objemu dat
- V blízké budoucnosti výměna hardwaru z důvodů nově nastupujících vyspělejších technologií
- Firma se nezbavuje přechodem na novou technologii o neustálou starost o data, zodpovědnost nepřechází na jiný subjekt

4.3 Externí zálohování

CLOUD

Zálohování s využitím CLOUD

V tomto případě bude **CLOUD** fungovat jen pro zálohování. V první fázi se nejedná o absolutní převedení celého systému externě mimo firmu, ta jen začne využívat nabídky ukládání záložních dat mimo své působiště. Ve firmě zůstane i celá infrastruktura, včetně zálohovacího serveru. Zálohovací program si v prostoru **NAS** vytvoří vlastní virtuální knihovnu (**VTL**). Z lokálního úložného prostoru, který tvoří zařízení **NAS**, je záloha replikována přes externí konektivitu do vzdáleného úložiště. To vyžaduje pořízení rychlého pevného optického spojení. Jelikož využíváme rezervní vlákno do datového centra stávajícího poskytovatele internetového připojení a nepotřebujeme plnou kapacitu linky směřovat do Internetu, je nová linka cenově dostupná.

V úložišti **NAS** probíhá redukce dat pomocí techniky **de-duplikace**. **De-duplikace** šetří prostor na cílové periférii, ale také významně redukuje zatížení komunikační linky do vzdáleného datového centra. V případě vyšších nároků na spolehlivost a dostupnost zálohovacího serveru **NAS** je možné zdvojit datovou cestu přes více než jeden **Ethernet** přepínač 10 Gb. Firma i nadále využívá původní program pro zálohování, pouze si v souladu s licenční politikou výrobce pořídí plug-in pro **VTL**. Pro správce **IT** infrastruktury se tak nic podstatného nemění, pouze místo fyzické robotické páskové knihovny začne používat knihovnu virtuální. Struktura zálohovacích úloh a jejich časování se mění jen minimálně, a to s ohledem na externí konektivitu.

Zálohování v prostředí CLOUD

Další fází je úplné převedení firemní infrastruktury do **CLOUD**. Firma si ponechá pouze některé speciální technologické servery a počítače, které nelze vyčlenit například z důvodu využívání některého ze starších druhů sběrnic (například sériová linka, jako komunikační rozhraní s terminálem docházky apod.). Dále bude nutné zvýšit spolehlivost a propustnost internetové komunikace. Ta je klíčovým požadavkem

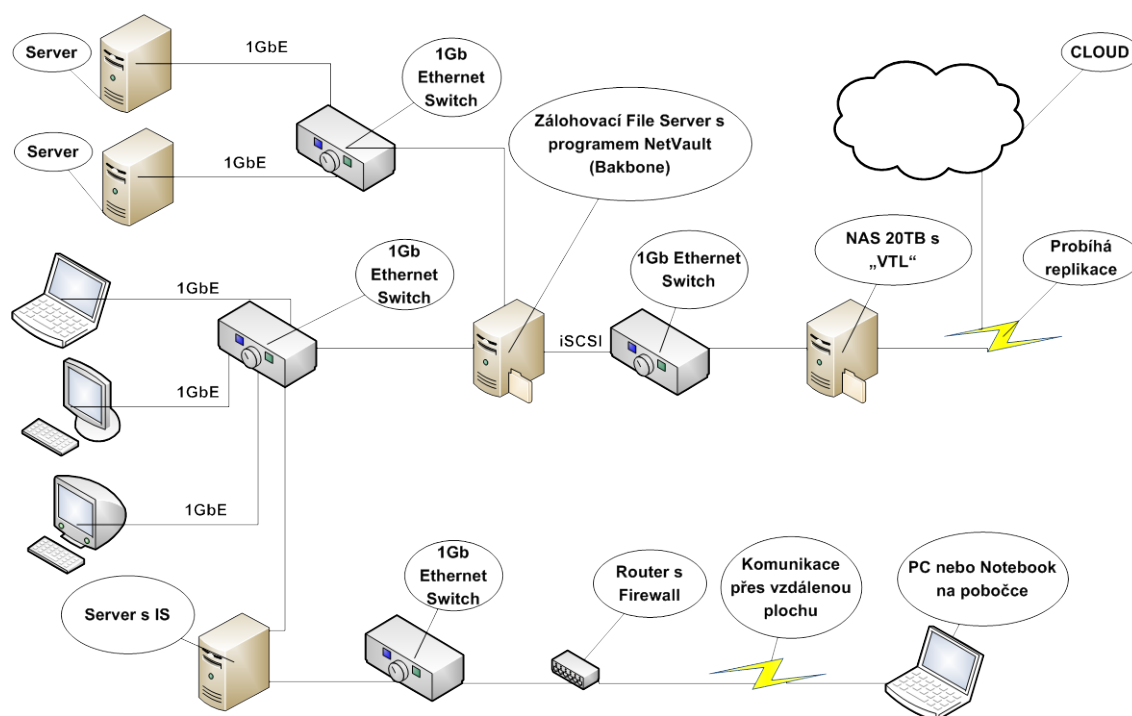
při plném využití řešení **CLOUD**. V **CLOUD** budou vytvořeny všechny servery, nezbytné pro chod **IT** infrastruktury firmy. Virtuální servery běží na požadovaných operačních systémech, mají definovaný výkon (počet procesorových jader a velikost operační paměti), mají k dispozici požadovaný virtuální úložný prostor. Na serverech jsou zkonfigurovány požadované databázové struktury.

Poskytovatel **CLOUD** nabízí různé varianty zabezpečení dat, od replikace zpět k uživateli, až po globální replikace mezi geograficky vzdálenými datovými centry, rozmístěnými na různých kontinentech.

Všechny parametry serverů, velikosti úložných prostorů, stupně replikací, operační systémy a databázové programy jsou **tarifikovány** (souborný pojem pro zjišťování hovorného v telefonních sítích) na měsíční bázi za jednotku výkonu, prostoru a času využití. Zálohování v infrastruktuře **CLOUD** neprobíhá tradiční metodou, ale výhradně násobením datových entit pomocí replikací. Tradiční koncept zálohování, jako vytváření bezpečnostních kopií na separátní média se tím od základu mění. Uživatel již nepotřebuje zálohovací server ani zálohovací program. Naopak musí dobře ovládat nástroje pro konfiguraci zdrojů a prostředků v **CLOUD**, včetně dokonalého monitorování stavu a včasného reagování na odchylky od normálu. V dobře postaveném **CLOUD** se praktici nemohou data ztratit vlivem havárie, mohou však být znehodnocena lidskou chybou nebo nekalým jednáním. O to je monitoring důležitější, než kdykoliv před tím. Pro konzervativní uživatele, kteří nedůvěřují bezchybné konektivě do Internetu, se nabízí varianta s replikacemi virtuálních serverů do vlastního virtuálního prostředí, které si vybudují doma mimo firmu nebo přímo ve firmě. Ovšem to záleží na managementu firmy. V případě výpadku konektivity se provoz přesměruje na lokální síť a po obnovení spojení se data opět vzájemně replikují do **CLOUD**. Toto řešení je ale samozřejmě zapláceno zvýšenými náklady na vlastní virtuální prostředí.

Volba způsobu zálohování tedy bude záležet na stupni virtualizace firemní **IT** infrastruktury a množství finančních prostředků, které bude chtít majitel do technologií investovat. Rozdíl nákladů mezi vlastní **IT** technikou a virtuálním prostředím v **CLOUD** je v Evropě ještě pro **CLOUD** nepříznivý. Dá se však přepokládat, že díky nastoupenému trendu se tento poměr v dohledné době obrátí.

Obrázek 6: Grafické schéma řešení CLOUD (Zdroj: autor a management společnosti)



Výhody:

Jedna z hlavních výhod je úspora investic do aplikačních a databázových serverů a datového úložiště v mateřské firmě a dalšího technického vybavení, včetně vybavení na pobočkách. V pobočných závodech zůstávají fyzicky **PC/NB**, které komunikují v terminálovém provozu s datovým centrem. Tím pádem odpadá povinnost zálohovat data jak v centrále, tak na pobočkách. Řešení **CLOUD** je flexibilnější, protože se jedná o službu lehce konfigurovatelnou a přizpůsobivou. Podle potřeby firmy je možné přikupovat větší místo pro data, měnit počet procesorových jader nebo objem operační paměti. Svou roli sehrává i **Thin Provisioning**, technologie přidělování prostředků podle okamžité potřeby a podle krátkodobých trendů čili mechanismus, který se uplatňuje v rozsáhlých centralizovaných storage systémech, **SAN** sítích, a storage virtualizačních systémech. Co se týká starostí spojených s úložištěm, lehčí to opravdu už být nemůže. Provozovatelé **CLOUD** řešení nabízí různé techniky tarifkace, které vyhovují rozličným požadavkům zákazníků. Někdy je výhodnější rezervovat určitý

výkon a kapacitu na pevně definovanou dobu, zatímco jindy může být výhodné nechat si tyto hodnoty přidělovat podle okamžité potřeby. Existuje například tarif pro datový prostor určitého objemu, který je celý k dispozici do té doby, dokud jej provozovatel **CLOUD** neposkytne jinému uživateli, který krátkodobě potřebuje okamžité navýšení. Trendy naznačují i značnou cenovou výhodnost. Jedná se o investici do budoucna, která se může bohatě vyplatit. První firmy mohou spustit lavinu. Kdo dřív přijde, ten dřív profituje. **CLOUD** je perspektivní krok směrem k větší pružnosti celého oboru IT a týká se nejen podnikové sféry, ale i běžných individuálních uživatelů. **CLOUD** velmi razantně proniká do oblasti chytrých telefonů, kde nabízí pronájem aplikací a pružný úložný prostor pro privátní nebo sdílené soubory (fotografie, hudba, kalendáře, kontakty atd.). **CLOUD** pro uživatele chytrých telefonů samozřejmě řeší i problematiku zálohování.

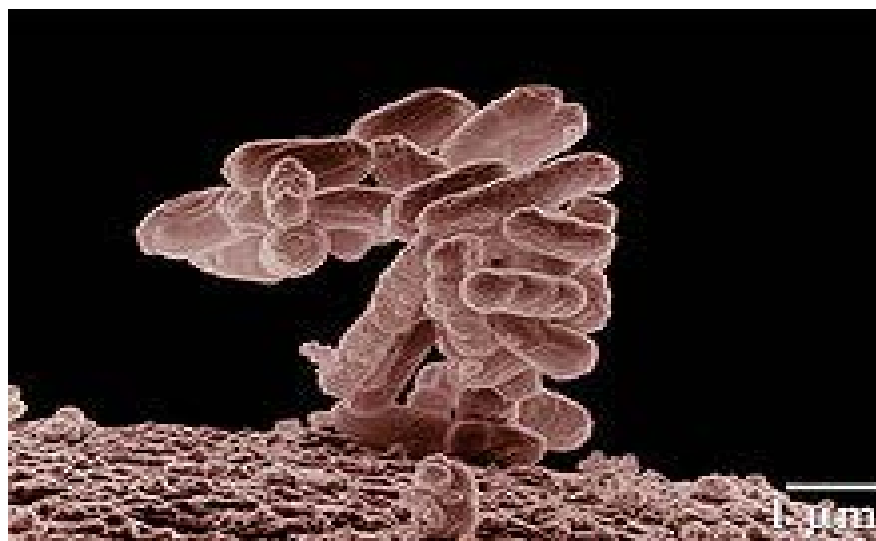
Nevýhody:

Protože se jedná o novou technologii, je na ni nahlíženo se značnou nedůvěrou a skepsí kvůli pocitu, že data nejsou v držení a správě firmy, ale v cizích rukou, což do značné míry svádí k myšlenkám, které se pojí se slovem zneužití. Z části se jedná o pravdivé pochybnosti, ale na druhou stranu jsou všechny přenosy kryptovány, tudíž dodavatel či poskytovatel služby garantuje v rámci **SLA** (Service Level Agreement) dohody, garanci přístupnosti a nezneužití. Teoreticky se může jednat i o podvodnou společnost, která se snaží získat citlivé informace, nebo jinak zneužít práva obsahu dat. Z právního hlediska se zdánlivé ohrožení privátních dat u cizího subjektu naopak může stát výhodným, protože **SLA** může pokrývat i pojištění rizika ztráty a zneužití informací, což v tradičním pojetí vlastní **IT** infrastruktury není možné. Firma musí počítat s průběžnými peněžními investicemi do udržovacích položek a potřebného navyšování úložného prostoru v **CLOUD**. Pro spolehlivé fungování **CLOUD** je kriticky naprosto nezbytná velmi kvalitní a výkonná konektivita. V neposlední řadě je těžké předpovídat, co se bude odehrávat v nedaleké budoucnosti a jestli technologii **CLOUD** za pár let někdo bude uznávat. Tím by se vložená investice obrátila v prach. **IT** oblast se neustále vyvíjí a s příchodem nové technologie se může tato problematika obrátit i jiným směrem.

4.4 Pohled do budoucnosti

V případě malého místa na pevném disku, není důvod propadat zoufalství. Studenti z čínské univerzity v Hongkongu přispěli s projektem biologického úložiště, kde se o data postará pozměněná **DNA bakterie E. Coli**, která ve vodě zpravidla indikuje fekální znečištění. Tento způsob není úplně nový, protože o biologickém uložení informací se mluví už jednu dekádu, hongkongskému týmu z tamní přírodovědecké fakulty se ale podařilo do **DNA** bakterie uložit i složitější binární data a struktury včetně obrázků. Po vložení informací žije upravená bakterie dál a později lze data z jejího nitra opět obnovit. Samozřejmě že o rychlostech zápisu a čtení se zde zatím hovořit nedá, nicméně pouhý jeden jediný gram **bakterií E. Coli** by postačil k archivaci 900 TB, tedy skoro 1 PB. Data mohou být dokonce i šifrovaná a zabezpečená biologickými ochrannými systémy bakterie, což je její imunitní systém.

Obrázek 7: Bakterie E. Coli (Zdroj: www.zive.cz/bleskovky/kolik-vazi-pevny-disk-s-kapacitou-900-tb-1-gram-a-je-zivy/sc-4-a-155546/default.aspx)



Jedná se o pouhý experiment. V té době mohou být na pořadu dne úplně nové a „mrtvé“ technologie včetně kvantových počítačů aj.⁵

⁵ Zive.cz [online]. 2011. 2011 [cit. 2011-05-30]. Kolik váží pevný disk. Dostupné z WWW: <<http://www.zive.cz/bleskovky/kolik-vazi-pevny-disk-s-kapacitou-900-tb-1-gram-a-je-zivy/sc-4-a-155546/default.aspx>>.

4.5 Doporučení jedné z navržených možností

Pro rozhodnutí, jestli jít moderně progresivní cestou nebo zůstat konzervativní, jsem podstoupil několikahodinové debaty s odborníky, jako je můj oponent, anebo kamarád, který se zabývá touto problematikou.

Dospěl jsem k názoru, že v dnešní nejisté době je dobré zůstat pevnýma nohama při zemi a přiklonit se k první variantě interního zálohování pomocí virtuálních knihoven, tedy že současný nevyhovující stav zálohování firemních dat nahradím technologií virtuálních knihoven, které jsou reprezentovány samostatným zařízením. Modelový typ Quantum DXi6520 je dodáván v rozsahu kapacit 8 - 32TB, s maximálně šesti 1GbE porty. Byla zvolena konfigurace s kapacitou 16TB, kterou lze do budoucna rozšiřovat po 8TB až na 32TB.

Koncepce virtuálních knihoven Quantum DXi dále nabízí zabudovanou podporu pro vzdálenou replikaci. Toho v našem případě bude využito pro přenos záložních dat mezi pobočnými závody a centrálou. Naopak na pobočkách budou replikována data z centrály a tím bude dosaženo ochrany dat proti zničení živlem v jednotlivých lokalitách. Virtuální knihovny DXi standardně využívají princip de-duplikace, snižující nároky na úložnou kapacitu a redukující požadavky na šířku komunikačních tras mezi vzdálenými knihovnami.

Kritéria, podle kterých byla dána přednost virtuálním knihovnám na bázi účelového samostatného zařízení, před širším využitím CLOUD na úschovu replikovaných záloh:

1) Nedůvěra zákazníků k úschově dat mimo vlastní firmu

Renomé tuzemských provozovatelů CLOUD řešení je zatím nízké. Obecně slabá vymahatelnost práva v našem regionu příliš neprospívá budování trvalých a důvěryhodných vztahů mezi firmami. Panuje obava ze zneužití citlivých informací ve prospěch nekalé soutěže. Tento stav lze částečně napravovat širokým nasazením silných kryptovacích technik, ale zbytečně to prodražuje celou koncepci CLOUD.

2) *Rozdílná legislativa u vyspělých provozovatelů CLOUD v USA a zákazníci v Evropě*

Vyspělé CLOUD zatím existují především v zámoří. Tuzemská nabídka služeb v CLOUD se začíná teprve rozvíjet. Využití zahraničního CLOUD naráží na rozdílné legislativy a případné právní spory nemají jednoznačně determinovanou podporu v tuzemsku. Nadějí pro budoucnost může být rozvoj evropské legislativy, která bude lépe hájit zájmy členských zemí v oblasti zpracování informací třetích stran.

3) *Dostatečně nerozvinutá konektivita s ohledem na výkon a spolehlivost*

Další parametr, který brzdí rychlejší rozšiřování internetových aplikací je rychlost samotné konektivity. Nejrychlejší sítě se u nás zatím mohou honosit vysoké školy a vědecké instituce. Rozšíření vysokorychlostního Internetu je na pořadu dne, ale zatím nedosahuje ani průměru dosahovaného například v rozvíjejících se a vyspělých asijských zemích. Levná, spolehlivá a rychlá konektivita do světové sítě je základním předpokladem rozvoje CLOUD v našem regionu.

4) *Úspora výdajů na zaškolení pracovníků IT*

Nasazením nové technologie autonomních virtuálních knihoven Quantum DXi se pro správce IT infrastruktury a konfiguraci zálohovacího programu nic podstatného nemění. Virtuální knihovna funguje prakticky stejně jako ta robotická, jen získáme vyšší rychlost, flexibilitu výkonu a kapacity. Využití replikace mezi knihovnami DXi je jen další bonus, spolu s úsporou kapacity díky de-duplikaci.

5) *Vazba na existující sadu záloh, uložených na páskových médiích LTO*

Virtuální knihovny DXi splňují i další důležitý požadavek, a to kompatibilitu s médii LTO, doposud používanými pro zálohování v robotické knihovně Scalar 100. Technologie DXi pracuje s virtuálními páskami, které mají stejný formát i kompresi, jako pásy fyzické. Fyzická média lze do DXi kdykoliv importovat nebo naopak virtuální média lze kdykoliv transformovat do fyzické podoby. K této funkci bude použito dosluhující robotické knihovny Scalar 100.

4.6 Přínos návrhu řešení

- *Úspora nákladů na IT specialisty*
- *Zvýšení bezpečnosti a spolehlivosti*
- *Úspora energie*
- *Ochrana proti živlu i na pobočkách*
- *Vyřešení nesourodosti médií*

VTL přivede nesčetnou řadu přínosů do firmy. Těmi největšími jsou ušetření místa, energie a lepší chladicí systémy. Určitě je dobré zmínit používání stejného zálohovacího programu, čímž odpadá nutnost zaškolování, jež se bude týkat jen nové technologie VTL. Používání de-duplikace, což je zvláštní technika komprese dat, jež dokáže zabránit výskytu stejného bloku dat na jednom úložišti. Příkladem může být, OS Windows 7, kdy z každého počítače není nutné mít uložená data ve VLT, ale jen souhrn rozdílů v datech. Velikost redukce je mnohdy v desítkách procent. Zálohování trvá jen několik desítek minut.

ZÁVĚR

Tato bakalářská práce klade důraz na nalezení optimálního řešení s ohledem na spokojenost zákazníka ve všech směrech. Hlavní bod se týká technických poznatků a náležitostí v IT oblasti zálohování dat a datových úložišť, které má firma zastaralé a nedostačující.

Při výběru z mnoha variant řešení jsem svou pozornost zaměřil na dva konkrétní, mou osobou vybrané, modelové návrhy. První model se týká implantování virtuální knihovny do stávajícího systému s možností replikace do vzdálených míst, protože virtuální knihovna funguje obdobně jako zastaralá robotická knihovna, ale je ve všech ohledech technicky precizněji konstruovaná vzhledem k požadavkům zákazníka. Druhý model je novinkou. U nás nepříliš běžný nebo známý – externí řešení zálohování CLOUD, které lze v krátkosti charakterizovat jako poskytování služeb či programů uložených na serverech na Internetu.

Po zvážení výhod a nevýhod daných variant jsem se rozhodl pro konzervativnější variantu virtuální knihovny, tedy že nahradím stávající nevyhovující systém technologií virtuálních knihoven, přesněji modelem typu Quantum DXi 6520.

Hlavní důvody, kterými bych rád podložil moje konečné rozhodnutí nad řešením CLOUD, se dají napsat v pěti heslovitých bodech. **Úspora nákladů na IT specialisty, zvýšení bezpečnosti a spolehlivosti, úspora energie, ochrana proti živlu i na pobočkách a vyřešení nesourodosti médií**

Jsem přesvědčen o vhodném výběru řešení pro firmu, které by mělo přispět ke kontinuální úspěšnosti a nijak nepřerušit či neovlivnit fungování firmy, protože s dobrou investicí přichází pozitivní výsledky, jenž nejsou vidět okamžitě, ale objeví se v budoucnu. Protože co může být lepší než pracovat v zjednodušeném, zrychleném a přehlednějším IT prostředí, které má usnadnit práci a tím i část života, který zde prožíváme.

SEZNAM LITERATURY

- 1) *Amazon* [online]. 2011. 2011 [cit. 2011-05-30]. Ec2. Dostupné z WWW: <<http://aws.amazon.com/ec2/>>.
- 2) BASL, J., BLAŽÍČEK, R. *Podnikové informační systémy*. Praha : Grada Publishing, 2007. 288 s. ISBN 978-80-247-2279-5.
- 3) *Cloud Computing* [online]. 2011. 2010 [cit. 2011-05-30]. Dostupné z WWW: <<http://www.cloud-computing.cz/>>.
- 4) DOSEDĚL, T. *Počítačová bezpečnost a ochrana dat*. 1. vyd. Brno : Computer Press, 2004. 190 s. ISBN 80-251-0106-1.
- 5) CHARVÁT, J. *Firemní strategie pro praxi*. Praha : Grada Publishing, 2006. 204 s. ISBN 80-247-1389-6.
- 6) NORTHCUTT, S. *Bezpečnost sítí*. CP Books, 2005. ISBN 80-251-0697-7.
- 7) O2 [online]. 2011. 2011 [cit. 2011-05-30]. Managed Backup and Restore. Dostupné z WWW: <http://www.o2.cz/corporate/198026-managed_services/114812-managed_backup_restore.html/>.
- 8) SODOMKA, P. *Informační systémy v podnikové praxi*. Brno : Computer Press, 2006. 352 s. ISBN 80-251-1200-4.
- 9) STANEK, W. R. *Microsoft Exchange 2000 Server Kapesní rádce administrátora*. 2001. 408 s. ISBN 8072264397.
- 10) TVRDÍKOVÁ, M. *Aplikace moderních informačních technologií v řízení firmy*. Praha : Grada Publishing, 2008. 176 s. ISBN 978-80-247-2728-8.

- 11) VOŘÍŠEK, J. *Strategické řízení informačního systému a systémová integrace*. Praha : Management Press, 2006. 324 s. ISBN 80-85943-40-9.
- 12) Zive.cz [online]. 2011. 2008 [cit. 2011-05-30]. Cloud Computing. Dostupné z WWW: <<http://www.zive.cz/Clanky/Cloud-computing-slibna-budoucnost-nebo-marketing/sc-3-a-144443/default.aspx/>>.
- 13) Zive.cz [online]. 2011. 2011 [cit. 2011-05-30]. Kolik váží pevný disk. Dostupné z WWW: <<http://www.zive.cz/bleskovky/kolik-vazi-pevny-disk-s-kapacitou-900-tb-1-gram-a-je-zivy/sc-4-a-155546/default.aspx/>>.

SEZNAM OBRÁZKŮ

Obrázek 1: Robotická knihovna ADIC SCALAR 1000 LTO (Zdroj: autor a management společnosti)	30
Obrázek 2: Robotická knihovna ADIC SCALAR 100 LTO (Zdroj: autor a management společnosti)	31
Obrázek 3: Série zařízení Quantum Storage DXi s mimo jiné Quantum Dxi6520 (Zdroj: www.peripheralstorage.com/deduplication/quantum/quantum-dxi.htm).....	42
Obrázek 4: Série zařízení poslední generace od EMC Data Domain (Zdroj: www.altimate.pt/local/1b57de9018c00c17/Newsletter.html)	42
Obrázek 5: Grafické schéma řešení virtuální knihovny (VTL) (Zdroj: autor a management společnosti)	43
Obrázek 6: Grafické schéma řešení CLOUD (Zdroj: autor a management společnosti)	48
Obrázek 7: Bakterie E. Coli (Zdroj: www.zive.cz/bleskovky/kolik-vazi-pevny-disk-s-kapacitou-900-tb-1-gram-a-je-zivy/sc-4-a-155546/default.aspx)	50